The Influence of Vocabulary Size on the Organization of the Second-Language Mental Lexicon

Bachelor thesis
Utrecht University
English Language and Culture

Jelle van der Werff
4187288

Supervisor: dr. Koen Sebregts
Second reader: dr. Elena Tribushinina

Abstract

An experiment was conducted investigating the influence of L2 speakers' vocabulary size on phonological and semantic priming effects in a visual lexical decision task. The aim was to investigate whether the relative effect of phonological priming, compared to semantic priming, would be larger for speakers with smaller vocabularies, as these speakers have fewer semantic connections between words in their mental lexicon. Before the lexical decision task, a LexTALE test for measuring L2 vocabulary size was carried out. The study's results suggest that phonological priming effects are stronger for speakers with larger L2 vocabularies. Furthermore, the results suggest that semantic priming effects are stronger than phonological priming effects for speakers with smaller L2 vocabularies. Different explanations for the findings are discussed.

Keywords: *Mental lexicon, lexical decision task, priming, semantic priming, phonological priming, L2, second-language, vocabulary size*

Table of Contents

1. Introduction

It is widely assumed that words are stored in the human mind in a specific 'mental lexicon' (ML) (e.g. Aitchison, 2012, p. 11). A major question in the field of psycholinguistics remains how words are then organized in this ML. In recent years, more research has focused on the organization of second-language (L2) MLs (e.g. Trofimovich & Mcdonough, 2011). In order to add to the existing body of research and increase our knowledge of the structure of the L2 ML, the present study will investigate what influence the size of the L2 learner's vocabulary has on the nature of the connections between words in the speaker's L2 ML. In order to investigate this, a small experiment will be conducted using the lexical-decision task (LDT) paradigm. The remainder of this chapter introduces the ML, and an overview is provided for the different types of studies into the organization of the ML and these studies' results. Subsequently, the L2 ML will be discussed and this study's research question will be introduced.

1.1. The mental lexicon

The ML is proposed as a sort of mental dictionary where words are stored and from which words are retrieved quickly for use in speech (e.g. Aitchison, 2012, pp. 4-16). Different motivations exist for assuming that words are not stored randomly in the speaker's mind, and that careful organization of those words in something like a mental dictionary is required. Firstly, an average person's vocabulary is very large. An average educated speaker of English is assumed to know at least 50,000 words (Aitchison, 2012, p. 7). Secondly, word retrieval is incredibly fast. Experimental evidence has shown that listeners are able to recognize words in 200ms or less from the word's onset (e.g. Marslen-Wilson & Tyler, 1980). Lastly, evidence from different types of studies have given an indication as to how this ML might be organized, and have shown that words are, in fact, carefully organized. A selection of these studies and their methods will be discussed below.

1.1.1. Speech errors

One way of studying the organization of the ML involves studying speech errors. Recorded slips-of-the-tongue (SOTs) are assumed to be indicative of how words are organized in the ML. Broadly, two types of SOTs can be defined: selection errors and assemblage errors (Aitchison, 2012, p. 21). Selection errors include speech errors such as saying *yesterday* instead of *tomorrow* or *beggars* instead of *burglars.* Assemblage errors include errors such as saying *patter-killer* instead of *caterpillar* (Aitchison, 2012, p. 21). Now, as assemblage errors seem to be solely phonological mishaps, they do not reveal much about an organization of words in the ML in which semantic connections between words play an important role. Therefore, the most interesting SOTs for studying the ML are selection errors. In these errors, the intended word (target) is replaced by a word that is either semantically or phonologically related (the error). As Dell & Reich note, “phonological and semantic similarity jointly affect the chances that a given word will substitute for an intended word” (1981, p. 626). This last observation already reveals something about how the ML might be organized, namely that connections between words can either be phonological or semantic in nature. Therefore, in the ML, a word such as *house* will have connections to both semantically related words such as *garden* and *bedroom* and to phonologically related words such as *mouse* or *hose*. Although speech errors provide interesting information for investigating the ML, SOT data is relatively hard to obtain due to their unpredictable occurrence. Corpora of speech errors have been established for providing easy access to these data, one of the most famous being Fromkin’s speech error database, which contains a collection of SOTs collected by the author from her own speech or that of her friends and colleagues (Fromkin, 1973). Corpora such as Fromkin’s are useful for studying the ML. However, as Aitchison explains, they should be used with care (2012, p. 23). Both the writing down and the interpretation of the SOTs are subjective and thus likely to contain errors. However, other types of evidence are, of course, available and will be discussed below.

1.1.2. Neurological and neuropsychological research

Ever since the areas of Broca and Wernicke were found to be associated with language comprehension and production (Broca, 1861; Wernicke, 1874), neurological research has proven valuable for the study of language. The organization of words in the ML has also been a topic of study within this discipline. For instance, an experiment conducted on brain-damaged patients showed problems with the production of words from specific semantic categories, depending on which areas of the brain were affected (Damasio et al., 1996). This finding has led the researchers to conclude that physical and contextual characteristics of the object denoted by the word influence how and where that word is stored (p. 504). Similarly, evidence from an electroencephalogram (EEG) study measuring the event-related potential (ERP) known as N400, which is associated with semantic integration (e.g. Van Berkum, Hagoort, & Brown, 1999), has found that when a sentence given to the participants ended with a word from an unexpected semantic category, N400 reactions were greater than when the sentence ended with a word from an expected semantic category, despite both words being logically impossible in the given context (Federmeier & Kutas, 1999). Another experiment using EEG showed that different locations for the storage of content and function words can be identified in the brain (Brown, Hagoort, & Ter Keurs, 1999). These experiments provide further evidence for an organization of the ML where words are stored more closely together according to their semantic relatedness. In Brown, Hagoort, & Ter Keurs's experiment (1999), this meant little physical distance between words of the same type. However, in general, words that are semantically related will be closer together on a more abstract level, and might not be stored close together physically.

1.1.3. Experimental psycholinguistics

The last type of evidence discussed here comes from experimental psychology. A number of paradigms exist for studying word organization in the ML, three of which will be dealt with here. In the 1980s, a paradigm known as gating became popular in linguistic research. In a gating experiment, increasingly long segments of a spoken word are presented (Grosjean, 1980). After each sound segment the participant is then asked to identify the word they hear (Grosjean, 1980, p. 270). The gating paradigm has, for instance, been used to demonstrate the effect of word frequency and phonological similarity on word recognition (Metsala, 1997). Here, it was shown that a high-frequency word with few phonological neighbours is recognized fastest, thus strengthening the view that connections to words in the ML become stronger as a word is used more frequently (p. 54). Another paradigm used in psycholinguistic research for investigating lexical access is phoneme monitoring. In a phoneme monitoring task, the participant is asked to find a specific sound, say [b], in a spoken sentence. The time it takes for the participant to find the sound is assumed to be indicative of the time it takes to process the word that precedes the sound (Aitchison, 2012, p. 28). The last paradigm that we will discuss here and which will be adopted in the experimental part of the present study is the lexical decision task (LDT) (Aitchison, 2012, p. 27). In a LDT both words and non-words are presented to the participant either through sounds or visual display. The participant is asked to decide whether a presented ‘word’ is a word or a non-word. During the experiment, the participants’ reaction times (RTs) are measured. These RTs are then interpreted to give information about the difficulty of processing this word or non-word. LDTs are often used in combination with priming. In a LDT with priming, a word (the target) is preceded by a word (the prime) that is either related or unrelated to the target. For instance, the word *goal* might be preceded by the word *football* (related) or by the word *coffee* (unrelated). RTs for targets with related primes are then compared to the RTs for targets with unrelated primes. In general, RTs for targets with related primes are shorter than RTs for targets with unrelated primes (e.g. Neely, Keefe, & Ross, 1989). The connections between the related primes and their targets can be semantic or phonological in nature (e.g. Lukatela, Frost, & Turvey, 1998). The LDT used in the present study will be elaborated further in the Method section below.

To conclude, many different types of evidence are available for researchers investigating the way words are stored in the ML. On the basis of this research, different models for the organization of the ML and for accessing the lexical information in the ML have been proposed, such as Collins & Loftus's extended Spreading Activation Model (1975) and Dell's interactionist model for Spreading Activation (1986). While it is beyond the scope of the present study to discuss these models in much detail, some general remarks on how the ML might be organized will now be made.

1.2. The organization of words in the mental lexicon

Lexical entries in the ML are generally thought to exist on a number of different levels (Radford, 2009, p. 205). A distinction is made between the lemma (Kempen & Huijbers, 1983), which contains information about the entry's meaning or content, and form, which contains information about the entry's morphological and phonological properties (Radford, 2009, p. 205). Evidence for this comes, for instance, from the difference between the English word *scissors* and the German equivalent *Schere*. While on the level of the lemma both concepts will have the same meaning, on the level of their form, the words differ in that the English word is always a plural, while the German word is singular (Radford, 2009, p. 205). A widely assumed property of the ML is that when lemmas are activated, that activation 'spreads' along the connections between related lemmas (Collins & Loftus, 1975). For instance, when the concept of *car* is activated, related concepts such as *truck* or *bus* become slightly activated as well. This idea is known in the literature as *spreading activation* (Collins

& Loftus, 1975). Spreading activation can be demonstrated through priming experiments (e.g. Collins & Loftus, 1975). As mentioned above, these experiments activate a concept, and subsequent activation of similar concepts is then measured. Spreading activation has been demonstrated repeatedly, and the existence of such a mechanism is assumed in the rest of this paper.

1.3. The second-language ML

As the present study will investigate priming effects in the second-language (L2) learner's ML, it is relevant to discuss what is known about the differences between the first-language (L1) ML and the L2 ML.

Research into the structure of the L2 ML is relatively new, and contradicting theories exist. For instance, Channell concludes that "[e]vidence that the L2 user's mental lexicon of a given language resembles the L1 user's lexicon is sparse" (1990, p. 29), while later research such as that by Wolter concludes that the L2 ML might not be "[. . .] nearly as randomly and loosely structured as past research seem[s] to indicate" (2001, p. 66). Results from word association (WA) tasks, in which the participant is asked to respond to a presented word with the first associated word that comes to the participant's mind, indicate similarities between L2 speakers' responses and responses from native speaker children (Wolter, 2001, p. 42). In studies involving WA tests, the relationship between the 'prompt word' and the participant's answer are commonly described as being of one of three types: (1) syntagmatic, (2) paradigmatic, or (3) phonological (Zareva, p. 129). Syntagmatic relations between words are, for instance, synonyms, antonyms etc. Paradigmatic relations are, for instance, collocations and adjective–noun pairs. Lastly, phonological relations involve two words with phonological similarities. Previous WA research has concluded that knowledge about individual words seems to negatively influence the number of responses with phonological similarity to the prompt word (Wolter, p. 65). The research further concludes that this shift from responses with phonological connections to the prompt word to more responses with paradigmatic–syntagmatic relations to the prompt word only happens for moderately well-known words, and not for well-known words (p. 66).

To conclude, competing theories on the similarity between the L1 and L2 ML exist. Furthermore, the influence of phonological and semantic relations on the organization of words in the L2 ML is yet to be fully understood. Moreover, the influence of vocabulary size on these relations is unclear. This research will aim to answer these questions with experimental evidence.

1.4. Research question, expectations, and assumptions

This study's research attempts to answer the question of whether the relative effect of semantic priming, when compared to phonological priming, becomes stronger as the L2 learner's vocabulary grows in size. The expectation is that that for speakers with smaller L2 vocabularies, phonological priming effects will be stronger than semantic priming effects, and that for speakers with larger L2 vocabularies, semantic priming effects will be stronger than phonological priming effects. This is expected as speakers with larger vocabularies will have more semantic connections between words in the ML, facilitating the effect of semantic priming. This expectation is further based around the conclusion drawn by Wolter that "it appears that although the phonological connections between words in the L2 ML do sometimes seem to take precedence over semantic connections for words that are moderately well known, this phenomenon tends to fade as greater understanding of individual words is gained [. . .]" (2001, p. 66). As speakers' knowledge of individual words is relatively hard to quantify, it is assumed in this study that general word knowledge, or, more specifically, vocabulary size positively correlates with individual word knowledge.

2. Method

In the present experiment, the visual LDT paradigm was used to investigate the effect of vocabulary size on semantic and phonological priming effects in L2 speakers of English. The experiment was based on the original visual LDT by Rubenstein, Garfield, & Millikan (1970) and on Meyer & Schvaneveldt's visual LDT with priming (1971). Furthermore, the LexTALE test (Lemhöfer & Broersma, 2011) was used for measuring the participants' vocabulary size.

2.1. Participants

Twenty Dutch L2 speakers of English volunteered for participation in the experiment (13 women and 7 men). The participants were selected to have as widely differing levels of English proficiency as possible. All participants were between 20 and 50 years of age (M = 27.1, SD = 8.52). Participants were randomly assigned to either group 1, which was the semantic priming condition (n = 10) or to group 2, the phonological priming condition (n = 10).

2.2. Materials

A Macbook Pro (model mid-2012) running Ubuntu 10.10 was used to carry out the experiment. The LDT experiment was designed using the ZEP Experimental Control Application (Veenker, 2013). RTs were recorded using a BeexyBox for millisecond precision. Before the experiment, a LexTALE English proficiency test (Lemhöfer & Broersma, 2011) was administered. The LexTALE test was downloaded from the creators' website and subsequently loaded into MatLAB R2015aSP1.

2.3. Stimuli

Both conditions of the LDT contained 40 prime–target pairs, 40 unrelated word–target pairs and 40 pseudoword–target pairs. The 40 prime–target pairs were different for both conditions. In the semantic priming condition, the primes were semantically related to the target, whereas in the phonological priming condition, the primes were phonologically related to the target. The target words were the same 40 words for both conditions, and also within the conditions for the prime–target pairs and the unrelated word–target pairs. An example is given in Table 2.1.

Table 2.1
Example of Stimuli for Both Conditions

	Semantic priming condition	Phonological priming condition
Related prime–target pair	*black – white*	*height – white*
Unrelated prime–target pair	*serious – white*	*serious – white*
Pseudoword–target pair	*adair – educational*	*adair – educational*

The word pairs were randomly presented to the participants. The complete list of stimuli can be found in the Appendix. To ensure that the participants were familiar with the words presented as stimuli all English words were among the top 3000 most frequent words in the Corpus of Contemporary American English, as available online (Davies, 2016). All stimuli had a length between 2 and 12 letters. The semantic primes were partly taken from Nelson, McEvoy & Schreiber (1998), which contains results from free association experiments. Phonological primes were then selected to have as little semantic similarity to the target words as possible. Pseudoword primes were taken from the Plausible Non-Words list by Paul Meara et al. (Meara, 2016). Additional pseudoword primes were taken from Elgort (2011).

2.4. Procedure

Participants were placed in a comfortable and quiet room, mostly in their own homes, behind a desk with the Macbook Pro placed directly in front of them. The experiment was not held in the same room for each participant, due to time limitations of the present study. However, the rooms in which the participants were tested were chosen so as to contain as few distracting elements as possible. Before the actual experiment, the LexTALE test for measuring English proficiency (Lemhöfer & Broersma, 2011) was administered. Beforehand, participants were instructed to take as much time as they needed to answer the questions in the LexTALE test. After completion, the LDT was started. First, the participants received instructions from the experimenter on how to answer the questions in the LDT, and on how to use the BeexyBox. In the instructions, emphasis was placed on answering the questions quickly and accurately. After four practice stimuli, the participants had the opportunity to ask the experimenter questions on the experiment. Subsequently, the experiment started. In the experiment, words were presented one at a time. The words were presented in pairs, one after the other, and these pairs were of one of the three types described above. For participants in group 1, the words in the 'related' pairs had a semantic connection. For participants in group 2, the words in the 'related' pairs had a phonological connection. The participants' task was to decide whether both words from the presented pair were existing English words or not. Participants could respond after the second word was presented. If both words existed, participants were asked to press the 'yes' button on the BeexyBox. If one or both words were non-existing words, participants were asked to press the 'no' button on the BeexyBox. The procedure of the LDT is illustrated in Figure 2.1.

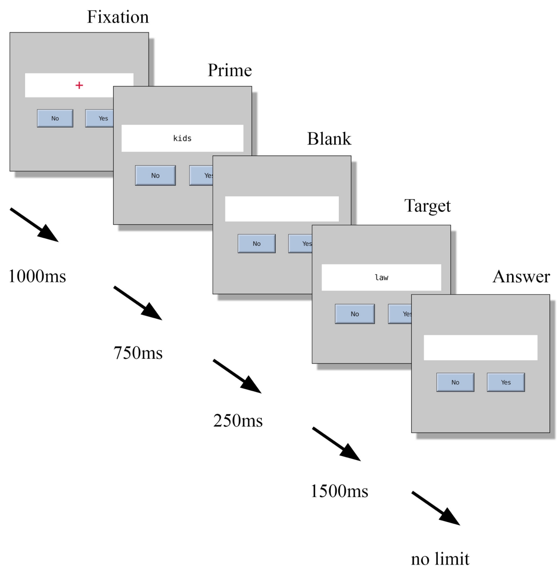

Figure 2.1. Procedure of the Experiment.

Primes were displayed for 750ms. This was decided after running several pilot experiments, as the initially chosen stimulus duration of 500ms proved too short for some participants to interpret the stimuli. Between the first and second word of the pair, a blank screen was displayed for 250ms. The target word was then displayed for 1500ms. As the experiment was self-paced, the participants could then respond at any point after presentation of the second stimulus. RTs were recorded from the onset of the second stimulus. Making the experiment self-paced was decided after pilot runs of the experiment, where participants indicated being confused by the presentation of the following stimulus, without having given an answer on the previous stimulus. Despite the fact that shorter stimulus onset asynchronies (SOAs) – the time between the onset of the first stimulus and the onset of the second stimulus – are more common (for a review, see De Groot, 1984), this SOA was chosen to ensure all participants were able to interpret the presented stimuli.

3. Results

For the analysis, all RTs for pseudoword–real word pairs were removed from the dataset, as these results were irrelevant to the rest of the analysis. Incorrect answers were removed from the dataset as well. Furthermore, RTs above 2000ms were removed, as these were obvious outliers. In line with observations by Baayen & Milin (2010, p. 4), a log transformation was performed on all RTs for enhanced normality. Subsequently, another 11 obvious outliers were removed based on the log RTs. For the remaining observations, means, standard deviations, and frequencies before the log transformation can be found in Table 3.1 below.

Table 3.1

Reaction Time Means, Standard Deviations and Frequencies (N) by Condition and Type of Word Pair

	Semantic priming condition			Phonological priming condition		
Type of word pair	Mean (ms)	Std deviation (ms)	N	Mean (ms)	Std deviation (ms)	N
Related prime–target	477	122	393	531	164	387
Unrelated prime–target	497	135	384	525	152	379

As is visible from the table, in the semantic priming condition, RTs were shorter for word pairs with a semantic connection than for word pairs that had no connection. Conversely, in the phonological priming condition, RTs were shorter for word pairs with no phonological connection than for word pairs with a phonological connection. However, these differences are very minimal, and unlikely to be above chance levels. To investigate these differences further, and to investigate the influence of English proficiency on the effect of semantic and phonological priming, a linear mixed effects model was built. This was done using the lme4 package (Bates, Mächler, Bolker, & Walker, 2009) within the R environment, version 3.3.2 (R Core Team, 2016). A number of models with different random and fixed effects were tried and compared for fitness. As factors such as ‘age’, ‘sex’, and ‘handedness’ did not positively influence the model's fitness, these were left out of the model. Fixed effects that were included in the model were 'type', 'condition', and 'LexTALE score'. A summary of the model can be found in Table 3.2.

Table 3.2
Linear Mixed-Effects Regression Predicting Log Reaction Times

Random effects	Variance	Std deviation	N	
Participant	0.024	0.022	20	
Word pair	0.000	0.155	80	
Fixed effects	**Estimate**	**Std error**	**t**	**p**
Intercept	6.318	0.254	24.883	.000
Type (related/unrelated)	0.101	0.065	1.560	.119
Condition	0.180	0.417	0.430	.672
LexTALE score	-0.001	0.003	-0.791	.440
Type x condition	-0.260	0.109	-2.381	.017
Type x LexTALE score	0.000	0.000	-1.166	.244
Condition x LexTALE score	0.002	0.005	-0.303	.765
Type x condition x LexTALE score	0.003	0.001	2.226	.026

As can be seen in the table, no main effects on log RT were found for the type of word pair (related/unrelated), condition (phonological priming/semantic priming) and LexTALE score. However, an interaction effect was found between type of word pair and condition on log RT and a three-way interaction effect was found between type of word pair, condition, and LexTALE score on log RT. This three-way interaction effect is illustrated in Figure 3.1.

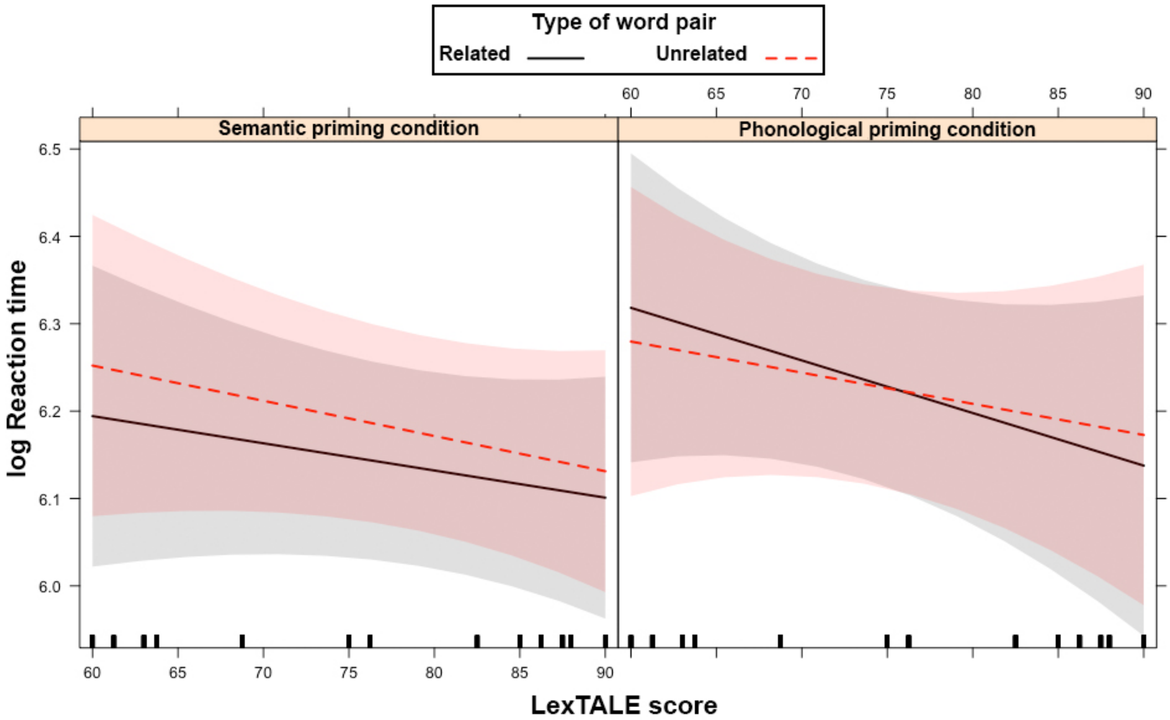

Figure 3.1. *Log Reaction Times with Confidence Intervals by Condition, LexTALE Score, and Type of Word Pair.*

The figure shows that in the semantic priming condition, pairs of related words gave shorter RTs for all LexTALE scores. However, in the phonological priming condition, pairs of related words only produced shorter RTs when the participants' LexTALE score was around 75 or higher, and not when the LexTALE score was lower. In fact, when looking at the RT lines for the phonological priming condition, one can observe higher RTs for related word pairs than for unrelated word pairs when LexTALE scores are below 75.

4. Discussion

4.1. Interpretation of the results

The present study investigated the influence of L2 vocabulary size on priming effects in a LDT. It was expected that the effect of semantic priming, as opposed to phonological priming, would become stronger as the L2 learner's vocabulary grew in size. However, the results suggest that this is not the case, and that, in fact, the opposite may be true. A three-way interaction effect on RTs was found between score on the LexTALE test, condition of the experiment (phonological priming/semantic priming) and type of word pair (related/unrelated). The results show no effect of phonological priming for speakers with the lowest LexTALE scores. However, phonological priming effects start to appear for participants with higher LexTALE scores.

Contrary to initial expectations, the experiment's results suggest that phonological priming effects become stronger as the L2 learner's vocabulary grows in size. As no main effects were found for 'condition', 'type', and 'LexTALE score' on RTs, and therefore no real priming effect was measured, it is difficult to draw conclusions about this. However, if these results are replicated in experiments with larger sample sizes, this could mean that in the developing ML, semantic connections between words are initially prioritized, and that phonological connections between words become stronger as the ML grows. This would contradict Wolter's observation that phonological connections between words in the ML take precedence over semantic connections for words that are moderately well known (2001, p. 66). However, Wolter's experiment used a WA task, and differences between his results and the present experiment's results might be explained by the difference in methodologies, as another explanation for the present study's results might be that relatively high L2 proficiency is required for interpreting the stimuli in a LDT phonologically rather than orthographically. Previous research has suggested that phonological similarity between primes and targets takes precedence over orthographic similarity (Lukatela, Frost, & Turvey, 1998). However, this has not yet been investigated for different levels of L2 proficiency. Therefore, a study investigating the difference between orthographic priming and phonological priming that uses L2 proficiency as a covariate, might shed more light on the present study's results. Still, interaction effects were found, suggesting that this study's methodology can prove valuable for investigating the L2 learner's ML.

4.2. Limitations and suggestions for future research

The present study has a number of limitations, mostly due to time considerations, which should be taken into account when interpreting the results, and future research should aim to eliminate at least some of these problems.

Firstly, a very small sample size was used (N = 20). While this was enough for the purposes of this preliminary study, more data is needed to be able to draw any serious conclusions from the data, and to confirm the effect of vocabulary size on the relative effect of semantic priming, when compared to the effect of phonological priming. Furthermore, this study's results show faster RTs for participants with higher LexTALE scores, despite the fact that no main effect was found for LexTALE score on RT. This might be due to the random effect of participant on RT. Furthermore, differences between different participants' LexTALE scores were relatively large ($M = 76.6$, $SD = 13.8$), which possibly influenced the results. A larger sample size is therefore needed to confirm the findings of this study.

Secondly, stimuli were hand-selected to meet a number of requirements, and future research should investigate whether these requirements were indeed met. For instance, minimal overlap in priming method is desired. This means that primes in the phonological condition should be selected to contain as little semantic similarity to the targets as possible, and vice versa. Furthermore, all stimuli should be selected to require a minimal level of L2 proficiency for interpretation as incorrect answers were now simply excluded from the analysis. Finally, actual priming effects for the used primes and targets should be investigated, for instance with the help of separate WA tasks, as no main priming effect was found.

Lastly, participants were not all tested in the same environment. Even though the rooms in which the participants were tested were selected to contains as few distracting elements as possible, distraction might still have influenced the results. Therefore, future research might require a more experimental setting.

5. Conclusion

This study investigated the influence of L2 vocabulary size on the relative effect of semantic and phonological priming, in order to learn more about the organization of the developing L2 ML. While the expectation was that semantic priming effects would become stronger than phonological priming effects as the L2 learner's vocabulary grows in size, the results suggest the reverse might be true. Namely, phonological priming effects are stronger for more advanced learners. If future experiments with larger sample sizes replicate the current study's findings, it may be hypothesized that in a developing L2 ML, semantic connections are prioritized and that phonological connections between words are developed later in the L2 acquisition process.

References

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. Chichester: Wiley-Blackwell.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Berkum, J. J. A. Van, Hagoort, P., & Brown, C. M. (1999). Semantic Integration in Sentences and Discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671.

Broca, P. (1861). Perte de la parole, ramollissement chronique et desstruction partielle du lob antérieur gauche de cerveau. *Bulletins de La Société d'Anthropologie*, *62*, 235–238.

Channell, J. M. (1990). Vocabulary acquisition and the mental lexicon. In J. Tomaszczyk & B. Lewandowska-Tomaszczyk (Eds.), *Meaning and Lexicography* (pp. 21–31). Amsterdam: John Benjamins Publishing Company.

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, 82(6), 407–428.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A Neural basis for Lexical Retrieval. *Nature*, *380*.

Davies, M. (2016). "Word frequency data". Retrieved from http://www.wordfrequency.info

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 283–321.

Dell, G. S., & Reich, P. A. (1981). Stages in Sentence Production: An Analysis of Speech Error Data. *Journal of Verbal Learning and Verbal Behaviour*. *20*, 611-629.

De Groot, A. M. B. (1984). Primed Lexical Decision: Combined Effects of the Proportion of Related Prime-Target Pairs and the Stimulus-Onset Asynchrony. *The Quarterly Journal of Experimental Psychology*, *36*(A), 253–280.

Elgort, I. (2010). Deliberate Learning and Vocabulary Acquisition in a Second Language. *Language Learning*, *61*(2), 367–413.

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing and Sentence Processing. *Journal of Memory and Language*, *41*.

Fromkin, V. A. (Ed.) (1973). *Speech errors as linguistic evidence* (Vol. 77). The Hague: Walter de Gruyter.

Grosjean, F. (1980). Spoken word recognition processes. *Perception & Psychophysics*, *28*(4), 267–283.

Lemhöfer, K., & Broersma, M. (2011). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325–343. http://doi.org/10.3758/s13428-011-0146-0

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological Priming by Masked Nonword Primes in the Lexical Decision Task. *Journal of Memory and Language*, *39*, 666–683.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition*, *25*(1), 47–56.

Meara, P. (2016). "PNWs". Retrieved from http:// http://www.lextutor.ca/freq/lists_download/pnwords.html

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in Recognizing Pairs of Words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic Priming in the Lexical Decision Task: Roles of Prospective Prime-Generated Expectancies and Retrospective Semantic Matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1003–1019.

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Radford A., Atkinson, M., Britain, D., Clahsen, H., & Spencer, A. (2009). *Linguistics: An introduction*. Cambridge: Cambridge University Press.

Trofimovich, P., & Mcdonough, K. (2011). *Applying priming methods to L2 learning, teaching and research: Insights from Psycholinguistics*. Amsterdam: John Benjamins Publishing Company.

Veenker, T.J.G. (2013). The Zep Experiment Control Application (Version 1.6.3) [Computer software]. *Utrecht Institute of Linguistics OTS, Utrecht University*. Available from http://www.hum.uu.nl/uilots/lab/zep/

Wernicke, C. (1874). Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis. Berlin: Springer Verlag.

Wolter, B. (2001). Comparing the L1 and L2 Mental Lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition*, *23*, 41–69.

Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research*, *23*(2), 123–153.

Appendix
List of All Stimuli Used in the Experiment

Semantic prime–target		*Phonological prime–target*	
Prime	*Target*	*Prime*	*Target*
theatre	play	week	weak
get	ready	beer	beard
computer	game	heat	beat
cook	kitchen	more	store
tail	cat	mood	food
car	truck	often	office
whisper	talk	cruise	crime
music	beat	ships	shoes
man	guy	page	paint
royal	queen	height	white
bottle	wine	me	tea
last	name	aunt	answer
running	shoes	part	parking
text	book	girl	guitar
turn	right	rest	ready
hair	beard	hat	cat
car	parking	glove	glass
black	white	same	name
question	answer	cold	could
prize	win	fame	game
lens	eye	wing	win
instrument	guitar	gray	great
buy	store	kids	kitchen
art	paint	mess	success
tie	suit	bone	boat
ship	boat	guide	guy
might	could	feed	need
desk	office	hook	book
steal	crime	cap	map
eat	food	raw	law
coffee	tea	trust	truck
small	great	stalk	talk
drink	glass	fruit	suit
want	need	mouse	house
world	map	might	right
garden	house	place	play
order	law	wings	winter
cold	winter	mine	wine
strong	weak	quarter	queen
fail	success	tie	eye

Unrelated word–target		*Pseudoword–target*	
Prime	*Target*	*Prime*	*Target*
circumstance	suit	activity	galpin
ignore	book	used	bance
magazine	shoes	nose	gumm
matter	boat	block	cambule
old	right	pants	nickling
live	need	these	berrow
sleep	house	choice	kiley
few	map	educational	adair
spread	glass	except	jarvis
sort	win	stick	buttle
itself	office	count	draconite
guess	queen	cigarette	moffat
resource	success	English	baldock
experience	guy	past	descript
system	wine	movie	bodelate
serious	tea	laugh	hapgood
him	store	image	gummer
castle	beard	faith	bastionate
hardly	law	extra	litholect
fear	ready	impossible	nonagrate
business	eye	weight	rudge
before	crime	file	aistrope
table	guitar	example	batcock
lamp	answer	professional	haque
rose	weak	here	balfour
audience	could	vegetable	escrotal
that	paint	start	charlett
civil	winter	song	contortal
indeed	white	travel	opie
meeting	name	metal	eckett
economy	truck	test	channing
orange	parking	tape	duffin
between	talk	either	eldred
line	food	salt	glandle
cite	game	prove	pauling
effort	play	leave	acklon
mean	kitchen	watch	mundy
card	beat	song	dowrick
off	cat	shut	connery
fuel	great	hang	degate